OIPE

ENTERED

**RAW SEQUENCE LISTING**
PATENT APPLICATION: **US/10/080,608A**

DATE: 03/21/2003 P.6
TIME: 13:34:55

Input Set : **A:\8471-010 b.txt**
Output Set: **N:\CRF4\03212003\J080608A.raw**

```
 3  <110> APPLICANT: Makowski, Lee
 4        Hyman, Paul
 5        Williams, Mark
 7  <120> TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
 9  <130> FILE REFERENCE: 8471-010-999
11  <140> CURRENT APPLICATION NUMBER: 10/080,608A
12  <141> CURRENT FILING DATE: 2002-02-21
14  <160> NUMBER OF SEQ ID NOS: 180
16  <170> SOFTWARE: FastSEQ for Windows Version 4.0
18  <210> SEQ ID NO: 1
19  <211> LENGTH: 28
20  <212> TYPE: PRT
21  <213> ORGANISM: Saccharomyces cerevisiae
23  <400> SEQUENCE: 1
24  Met Lys Gln Leu Glu Asp Lys Val Glu Glu Leu Leu Ser Lys Asn Tyr
25   1               5                   10                  15
26  His Leu Glu Asn Glu Val Ala Arg Leu Lys Lys Leu
27              20                  25
29  <210> SEQ ID NO: 2
30  <211> LENGTH: 28
31  <212> TYPE: PRT
32  <213> ORGANISM: Unknown
34  <220> FEATURE:
35  <223> OTHER INFORMATION: Probable variant of homo sapiens protein.
37  <400> SEQUENCE: 2
38  Thr Asp Thr Leu Gln Ala Glu Thr Asp Gln Leu Glu Asp Glu Lys Tyr
39   1               5                   10                  15
40  Ala Leu Gln Thr Glu Ile Ala Asn Leu Leu Lys Glu
41              20                  25
43  <210> SEQ ID NO: 3
44  <211> LENGTH: 28
45  <212> TYPE: PRT
46  <213> ORGANISM: Unknown
48  <220> FEATURE:
49  <223> OTHER INFORMATION: Probable variant of homo sapiens protein.
51  <400> SEQUENCE: 3
52  Ala Ala Arg Leu Glu Glu Lys Val Lys Thr Leu Lys Ala Gln Asn Tyr
53   1               5                   10                  15
54  Glu Leu Ala Ser Thr Ala Asn Met Leu Arg Glu Gln
55              20                  25
57  <210> SEQ ID NO: 4
58  <211> LENGTH: 35
59  <212> TYPE: PRT
```

RAW SEQUENCE LISTING                    DATE: 03/21/2003
PATENT APPLICATION:  **US/10/080,608A**        TIME: 13:34:55

Input Set : **A:\8471-010 b.txt**
Output Set: **N:\CRF4\03212003\J080608A.raw**

```
 60 <213> ORGANISM: Mus musculus
 62 <400> SEQUENCE: 4
 63 Val Leu Glu Thr Gln His Lys Asn Glu Arg Leu Thr Ala Glu Val Glu
 64 1               5                  10                  15
 65 Gln Leu Gln Lys Lys Leu Ser Thr Leu Ser Arg Glu Phe Lys Gln Leu
 66             20                  25                  30
 67 Arg Asn Leu
 68             35
 70 <210> SEQ ID NO: 5
 71 <211> LENGTH: 35
 72 <212> TYPE: PRT
 73 <213> ORGANISM: Homo sapiens
 75 <400> SEQUENCE: 5
 76 Cys Lys Glu Leu Thr Gly Glu Asn Glu Ala Leu Glu Lys Lys Ala Asp
 77 1               5                  10                  15
 78 Ser Leu Lys Glu Arg Ile Gln Tyr Leu Ala Lys Glu Ile Glu Glu Val
 79             20                  25                  30
 80 Lys Asp Leu
 81             35
 83 <210> SEQ ID NO: 6
 84 <211> LENGTH: 33
 85 <212> TYPE: PRT
 86 <213> ORGANISM: Homo sapiens
 88 <220> FEATURE:
 89 <221> NAME/KEY: VARIANT
 90 <222> LOCATION: 33
 91 <223> OTHER INFORMATION: Xaa = Any Amino Acid
 93 <400> SEQUENCE: 6
 94 Cys Gly Gly Val Gln Ala Glu Glu Gln Lys Leu Ile Ser Glu Glu Asp
 95 1               5                  10                  15
 96 Leu Leu Arg Lys Arg Arg Glu Gln Leu Lys His Lys Leu Glu Gln Leu
 97             20                  25                  30
```
**W--> 98 Xaa**
```
101 <210> SEQ ID NO: 7
102 <211> LENGTH: 33
103 <212> TYPE: PRT
104 <213> ORGANISM: Mus musculus
106 <220> FEATURE:
107 <221> NAME/KEY: VARIANT
108 <222> LOCATION: 33
109 <223> OTHER INFORMATION: Xaa = Any Amino Acid
111 <400> SEQUENCE: 7
112 Cys Gly Gly Met Arg Arg Lys Asn Asp Thr His Gln Gln Asp Ile Asp
113 1               5                  10                  15
114 Asp Leu Lys Arg Gln Asn Ala Leu Leu Glu Gln Gln Val Arg Ala Leu
115             20                  25                  30
```
**W--> 116 Xaa**
```
119 <210> SEQ ID NO: 8
120 <211> LENGTH: 32
```

RAW SEQUENCE LISTING                    DATE: 03/21/2003
PATENT APPLICATION:  US/10/080,608A     TIME: 13:34:55

Input Set : A:\8471-010 b.txt
Output Set: N:\CRF4\03212003\J080608A.raw

121 <212> TYPE: PRT
122 <213> ORGANISM: Unknown
124 <220> FEATURE:
125 <223> OTHER INFORMATION: Probable variant of Mus musculus protein
127 <400> SEQUENCE: 8
128 Val Lys Ser Leu Glu Asn Arg Val Ala Val Leu Glu Asn Gln Asn Lys
129 1               5                   10                  15
130 Thr Leu Ile Glu Glu Leu Lys Ala Leu Lys Asp Leu Tyr Ser His Lys
131             20                  25                  30
133 <210> SEQ ID NO: 9
134 <211> LENGTH: 33
135 <212> TYPE: PRT
136 <213> ORGANISM: Schizosaccharomyces pombe
138 <400> SEQUENCE: 9
139 Val Val Thr Leu Lys Glu Leu His Ser Ser Thr Thr Leu Glu Asn Asp
140 1               5                   10                  15
141 Gln Leu Arg Gln Lys Val Arg Gln Leu Glu Glu Glu Leu Arg Ile Leu
142             20                  25                  30
143 Lys
146 <210> SEQ ID NO: 10
147 <211> LENGTH: 885
148 <212> TYPE: PRT
149 <213> ORGANISM: Mus musculus
151 <400> SEQUENCE: 10
152 Met Glu Ile Gly Val Ser Val Ala Glu Cys Lys Ser Val Pro Gly Val
153 1               5                   10                  15
154 Thr Ser Thr Pro His Ser Lys Asp His Ser Ser Pro Phe Tyr Ser Pro
155             20                  25                  30
156 Ser His Asn Gly Leu Leu Ala Asp His His Glu Ser Leu Asp Asn Asp
157             35                  40                  45
158 Val Ala Arg Glu Ile Gln Tyr Leu Asp Glu Val Leu Glu Ala Asn Cys
159     50                  55                  60
160 Cys Asp Ser Ser Val Asp Gly Thr Tyr Asn Gly Ile Ser Ser Pro Glu
161 65                  70                  75                  80
162 Pro Gly Ala Ala Ile Leu Val Ser Ser Leu Gly Ser Pro Ala His Ser
163                 85                  90                  95
164 Val Thr Glu Ala Glu Pro Thr Glu Lys Ala Ser Gly Arg Gln Val Pro
165                 100                 105                 110
166 Pro His Ile Glu Leu Ser Arg Ile Pro Ser Asp Arg Met Ala Glu Gly
167             115                 120                 125
168 Glu Arg Ala Asn Gly His Ser Thr Asp Gln Pro Gln Asp Leu Leu Gly
169         130                 135                 140
170 Asn Ser Leu Gln Ala Pro Ala Ser Pro Ser Ser Ser Thr Ser Ser His
171 145                 150                 155                 160
172 Cys Ser Ser Arg Asp Gly Glu Phe Thr Leu Thr Thr Leu Lys Lys Glu
173                 165                 170                 175
174 Ala Lys Phe Glu Leu Arg Ala Phe His Glu Asp Lys Lys Pro Ser Lys
175             180                 185                 190
176 Leu Phe Glu Glu Asp Glu Arg Glu Lys Glu Gln Phe Cys Val Arg Lys

```
177          195                  200                  205
178 Val Arg Pro Ser Glu Glu Met Ile Glu Leu Glu Lys Glu Arg Arg Glu
179     210                  215                  220
180 Leu Ile Arg Ser Gln Ala Val Lys Lys Asn Pro Gly Ile Ala Ala Lys
181 225                  230                  235                  240
182 Trp Trp Asn Pro Pro Gln Glu Lys Thr Ile Glu Glu Gln Leu Asp Glu
183                  245                  250                  255
184 Glu His Leu Glu Ser His Arg Arg Tyr Lys Glu Arg Lys Glu Lys Arg
185              260                  265                  270
186 Ala Gln Gln Glu Gln Leu Gln Leu Gln Gln Gln Gln Gln Gln Gln Leu
187     275                  280                  285
188 Gln Gln Gln Gln Leu Gln Gln Gln Gln Leu Gln Gln Gln Gln Leu Gln
189     290                  295                  300
190 Gln Gln Leu Gln Gln Gln Gln Leu Ser Thr Ser Gln Pro Cys Thr Ala
191 305                  310                  315                  320
192 Pro Ala Ala His Lys His Leu Asp Gly Ile Glu His Thr Lys Glu Asp
193                  325                  330                  335
194 Val Val Thr Glu Gln Ile Asp Phe Ser Ala Ala Arg Lys Gln Phe Gln
195              340                  345                  350
196 Leu Met Glu Asn Ser Arg Gln Thr Leu Ala Lys Gly Gln Ser Thr Pro
197     355                  360                  365
198 Arg Leu Phe Ser Ile Lys Pro Tyr Tyr Lys Pro Leu Gly Ser Ile His
199     370                  375              380
200 Ser Asp Lys Pro Pro Thr Ile Leu Arg Pro Ala Thr Val Gly Gly Thr
201 385                  390                  395                  400
202 Leu Glu Asp Gly Gly Thr Gln Ala Ala Lys Glu Gln Lys Ala Pro Cys
203                  405                  410                  415
204 Val Ser Glu Ser Gln Ser Ala Gly Ala Gly Pro Ala Asn Ala Ala Thr
205              420                  425                  430
206 Gln Gly Lys Glu Gly Pro Tyr Ser Glu Pro Ser Lys Arg Gly Pro Leu
207     435                  440                  445
208 Ser Lys Leu Trp Ala Glu Asp Gly Glu Phe Thr Ser Ala Arg Ala Val
209     450                  455                  460
210 Leu Thr Val Val Lys Asp Glu Asp His Gly Ile Leu Asp Gln Phe Ser
211 465                  470                  475                  480
212 Arg Ser Val Asn Val Ser Leu Thr Gln Glu Glu Leu Asp Ser Gly Leu
213              485                  490                  495
214 Asp Glu Leu Ser Val Arg Ser Gln Asp Thr Thr Val Leu Glu Thr Leu
215     500                  505                  510
216 Ser Asn Asp Phe Ser Met Asp Asn Ile Ser Asp Ser Gly Ala Ser Asn
217     515                  520                  525
218 Glu Thr Thr Ser Ala Leu Gln Glu Asn Ser Leu Ala Asp Phe Ser Leu
219     530                  535                  540
220 Pro Gln Thr Pro Gln Thr Asp Asn Pro Ser Glu Gly Arg Glu Gly Val
221 545                  550                  555                  560
222 Ser Lys Ser Phe Ser Asp His Gly Phe Tyr Ser Pro Ser Ser Thr Leu
223              565                  570                  575
224 Gly Asp Ser Pro Ser Val Asp Asp Pro Leu Glu Tyr Gln Ala Gly Leu
225              580                  585                  590
```

RAW SEQUENCE LISTING                    DATE: 03/21/2003
PATENT APPLICATION:  US/10/080,608A     TIME: 13:34:55

Input Set  : A:\8471-010 b.txt
Output Set: N:\CRF4\03212003\J080608A.raw

```
226 Leu Val Gln Asn Ala Ile Gln Gln Ala Ile Ala Glu Gln Val Asp Lys
227         595             600             605
228 Ala Glu Ala His Thr Ser Lys Glu Gly Ser Glu Gln Gln Glu Pro Glu
229     610             615             620
230 Ala Thr Val Glu Glu Ala Gly Ser Gln Thr Pro Gly Ser Glu Lys Pro
231 625             630             635                     640
232 Gln Gly Met Phe Ala Pro Pro Gln Val Ser Ser Pro Val Gln Glu Lys
233             645             650             655
234 Arg Asp Ile Leu Pro Lys Asn Leu Pro Ala Glu Asp Arg Ala Leu Arg
235             660             665             670
236 Glu Lys Gly Pro Ser Gln Pro Pro Thr Ala Ala Gln Pro Ser Gly Pro
237         675             680             685
238 Val Asn Met Glu Glu Thr Arg Pro Glu Gly Gly Tyr Phe Ser Lys Tyr
239     690             695             700
240 Ser Glu Ala Ala Glu Leu Arg Ser Thr Ala Ser Leu Leu Ala Thr Gln
241 705             710             715                     720
242 Glu Ser Asp Val Met Val Gly Pro Phe Lys Leu Arg Ser Arg Lys Gln
243             725             730             735
244 Arg Thr Leu Ser Met Ile Glu Glu Glu Ile Arg Ala Ala Gln Glu Arg
245             740             745             750
246 Glu Glu Glu Leu Lys Arg Gln Arg Gln Val Arg Gln Ser Thr Pro Ser
247         755             760             765
248 Pro Arg Ala Lys Asn Ala Pro Ser Leu Pro Ser Arg Thr Thr Cys Tyr
249     770             775             780
250 Lys Thr Ala Pro Gly Lys Ile Glu Lys Val Lys Pro Pro Pro Ser Pro
251 785             790             795                     800
252 Thr Thr Glu Gly Pro Ser Leu Gln Pro Asp Leu Ala Pro Glu Glu Ala
253             805             810             815
254 Ala Gly Thr Gln Arg Pro Lys Asn Leu Met Gln Thr Leu Met Glu Asp
255             820             825             830
256 Tyr Glu Thr His Lys Ser Lys Arg Arg Glu Arg Met Asp Asp Ser Ser
257         835             840             845
258 Tyr Thr Ser Lys Leu Leu Ser Cys Lys Val Thr Ser Glu Val Leu Glu
259     850             855             860
260 Ala Thr Arg Val Asn Arg Arg Lys Ser Ala Ser Gly Leu Ala Leu Gly
261 865             870             875                     880
262 Gly Arg Asp Leu Arg
263             885
265 <210> SEQ ID NO: 11
266 <211> LENGTH: 3878
267 <212> TYPE: PRT
268 <213> ORGANISM: Homo sapiens
270 <400> SEQUENCE: 11
271 Met Glu Asp Glu Glu Arg Gln Lys Lys Leu Glu Ala Gly Lys Ala Lys
272 1               5               10              15
273 Ile Glu Glu Leu Ser Leu Ala Phe Leu Val Arg Gln Leu Ala Gln Phe
274             20              25              30
275 Arg Gln Arg Lys Ala Gln Ser Asp Gly Gln Ser Pro Ser Lys Lys Gln
276         35              40              45
```

RAW SEQUENCE LISTING ERROR SUMMARY    DATE: 03/21/2003
PATENT APPLICATION:  **US/10/080,608A**    TIME: 13:34:57

Input Set : **A:\8471-010 b.txt**
Output Set: **N:\CRF4\03212003\J080608A.raw**

**Please Note:**
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:6; Xaa Pos. 33
Seq#:7; Xaa Pos. 33
Seq#:71; Xaa Pos. 14
Seq#:143; N Pos. 8,9,10,11
Seq#:145; N Pos. 8,9,10,11
Seq#:147; N Pos. 8,9,10,11
Seq#:149; N Pos. 8,9,10,11

**Invalid <213> Response:**
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:77,78,165,166,167,170,171,172,173,175,176,177,178,179,180

**VERIFICATION SUMMARY**                  DATE: 03/21/2003
PATENT APPLICATION:  **US/10/080,608A**        TIME: 13:34:57

Input Set : **A:\8471-010 b.txt**
Output Set: **N:\CRF4\03212003\J080608A.raw**

```
L:98 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:32
L:116 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:32
L:5452 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:71 after pos.:0
L:5502 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (76) SEQUENCE:
L:6313 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:143 after pos.:0
L:6343 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:145 after pos.:0
L:6373 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:147 after pos.:0
L:6403 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:149 after pos.:0
```